DOCUMENT RESUME

ED 066 015 EM 010 088

AUTHOR	Hoffman, David E.; Cohen, Elizabeth G.
TITLE	An Exploratory Study to Determine the Effects of Generalized Academic Performance Expectations Upon the Activity and Influence of Students Engaged in a Group Simulation Game.
PUB DATE	5 Apr 72
NOTE	43p.; Paper presented at the American Educational Research Association Annual Meeting (Chicago, Illinois, April 5, 1972)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Evaluation; *Expectation; *Junior High School Students; *Self Evaluation; *Simulation; *Teacher Attitudes
IDENTIFIERS	*Self Fulfilling Prophecy

ABSTRACT

Junior high school boys ranked each other on perceived academic ability (PAA) and social power with great consistency. Holding social power constant, four-man groups were formed with either two high and two medium PAA boys or two medium and two low PAA boys, and each group played a non-academic, co-operative simulation game. In each condition, the boys with relatively higher perceived academic ability were significantly more active than boys with relatively lower perceived academic ability. Thus, the finding that perceived academic ability has significant effects on a non-academic situation may have distressing implications for all tracked educational systems and for attempts to break out of tracking by innovative methods like simulation. (RH)

FILMED FROM BEST AVAILABLE COPY

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ED 066015

A. E. R. A.

National Meeting

April 5, 1972

Chicago, Illinois

An Exploratory Study to Determine the Effects of Generalized Academic Performance Expectations Upon the Activity and Influence of Students Engaged in a Group Simulation Game

David E. Hoffman
Southern Oregon College

Elizabeth G. Cohen
Stanford University

EM 010 088

An Exploratory Study to Determine the Effects of Generalized Academic Performance Expectations Upon the Activity and Influence of Students Engaged in a Group Simulation Game

A critical question for classroom research today is the understanding of the self-fulfilling prophecy. If it is true that some of the variance in achievement is produced by the expectations for performance held by teachers and students, educators are somehow defeating their own ends. If initial expectations can impair a student's motivation and persistence at new educational tasks, it is vital that we understand the sources of these expectations in order that we may be able to modify them. We must have a satisfying explanation for this process before recommending effective methods of short-circuiting the self-fulfilling prophecy. At the moment it is not at all clear whether manipulation should begin with teacher expectation, with student expectations or with other more general features of the classroom which are necessary conditions for the occurrence of this process.

Brookover was one of the first sociologists to become interested in the self-fulfilling prophecy occurring in students. He was able to show that a student's academic self-concept predicted his performance, holding ability level constant (1965). Rosenthal and Jacobson's study of the manipulation of teacher expectations, tried to show that if the teacher expected certain children to bloom

intellectually, these children would be more likely to show IQ score gains over time than children for whom there was no such expectation (1968). There has been some question of the reliability of these results, but a more detailed study of sequences of interaction between teachers and students with high and low perceived academic achievement by Good and Brophy (1970) reveals highly suggestive differences in the activity levels of perceived high and low achievers and in the differential response of teachers to children with different perceived achievement levels. Teachers consistently favored "highs" over "lows" in demanding and reinforcing quality of performance. The "highs" were more frequently praised when correct and less frequently criticized when incorrect or unable to respond. Teachers were more persistent in eliciting responses from "highs" than from "lows". They were more likely to supply the answer and call on another child when responding to the "lows". Furthermore "highs" initiated more work-related contacts and created more response opportunities for themselves.

The teacher's expectations for the student and the student's expectations for himself are not the only variables involved. Early research shows that the other students in the class show considerable agreement in ranking students on achievement status. Lippitt and Gold were able to determine that the most important characteristic by which

children discriminated between one another were: the affective dimension, expertness in school activities and coercability. Of the three, the affective dimension and expertness in school activities were the most significant (Gold, 1958; Lippitt & Gold, 1959). In these studies there is a high degree of agreement between the students in a class on where the fellow-students rank on school success and ability. It may be inferred from these studies that other students function as "significant others" in the formation of a student's expectations for his own performance.

This interplay of expectations and evaluations suggests that the phenomenon of the self-fulfilling prophecy in the classroom can be seen as an instance of the operation of the activation of general expectations described by the Theory of Status Characteristics and Expectation States (Berger, Cohen & Zelditch, 1966). If it can be shown that willingness to participate on a <u>new and different task</u> is a function of expectations stemming from one's perceived position on a rank order of academic ability, there is a gain of all the advantages of that theory's explanation of how and why initial expectations, based on status come to diffuse to new situations and produce the same rank ordering of prestige and power on the new situations as in previous situations.

THEORETICAL EXPLANATION

What is this theoretical explanation and why would its use provide us with a basis for "short circuiting" the self-fulfilling prophecy? In an extensive body of empirical tests of the theory, researchers have been able to show that Diffuse Status Characteristics and General Performance Characteristics are capable of generating general expectations for competence or incompetence which will spread to new situations, even when the new situation has no relevance whatsoever either to the status characteristic or to the previous evaluations of the person's performance. A Diffuse Status Characteristic is a distinction in social rank commonly agreed upon throughout a society such as race, social class or organizational rank. In Cohen's laboratory study of interracial work groups white junior high school boys were much more active and influential than black junior high school boys on an unfamiliar non-academic game. The boys had no basis for prior evaluation of each other aside from knowledge of each other's race. Nevertheless, the power and prestige order of the new problem-solving group repeated the status order of the outside society with great regularity (Cohen, 1971). This is just one example of the many studies demonstrating the spread of expectations from a status order to a new situation (Berger, Cohen & Zelditch, 1966).

Social status taken from the outer society is not the same concept as ranking on a General Performance Characteristic

in a classroom. Recent developments by researchers on Expectation States have shown that if the only thing one group member knows about another group member is that he has higher scores on two related tests of performance, this status order (based on prior performance) will generate expectations for competence or incompetence on new tasks, even though the new tasks bear no resemblance to the tasks where evaluations were made (Freese, 1970).

Ranking on two or more performance capabilities related to each other is referred to technically as a General Performance Characteristic (Freese and Cohen, 1971).

Freese asserts that expectations associated with a specific performance characteristic will determine an observable power and prestige order in a new task situation if the characteristic is directly related to at least one other specific performance characteristic. If the specific performance characteristic can not be related to another, the generalization process is believed less likely to occur. An example of a specific performance characteristic would be the ability to speak well in public, to read well or to have computational skill. The formal definition of a specific performance characteristic follows.

Definition 1. Specific Performance Characteristic (C):

> A specific performance characteristic C is any characteristic of an actor for which it is the case that, from p's point of view:

(1) there are at least two states, such that there is associated with each state a different probability of successful task performance, and

(2) the states of the characteristic are differentially evaluated

(3) the characteristic is instrumental to a set of tasks (Freese, 1970, p. 13).

If competence is assigned on two specific performance characteristics which imply each other, Freese and Cohen speak of the creation of a general performance characteristic.

Definition 2. General Performance Characteristic (G):

A general performance characteristic is any characteristic of an actor for which it is the case that: there are at least two states, such that the states of the characteristic consist of a balanced set of symetrically related states of specific performance characteristics (Freese and Cohen, 1971. p. 6).

Freese and Cohen define some of the above terms as follows:

By set is meant a collection of at least two or more states of charcteristics. By balanced set is meant that all members of the set have the same evaluation, either positive or negative. By symmetrically related, or symmetrically relevant, is meant that possession of any one member of the set implies the expected possession from p's point of view, of all other members of the set (Freese and Cohen, 1971).

An example of a general performance characteristic would be to be a good athlete, to be a good student or to be recognized as good at mechanics. For this investigation, the perceptions class members hold of one another on Perceived Academic Ability was identified as an instance of a General Performance Characteristic.

The purpose of this study is to show that Perceived Academic Ability is capable of generating expectations in the same manner as a General Performance Characteristic has been shown to generate expectations in a laboratory setting. In order to test this proposition we had to find a task, outside the ordinary classroom setting which met the specifications of the theory for a situation where general expectations for competence stemming from a General Performance Characteristic would be activated. The task had to be interdependent, emotionally involving and unrelated to specific classroom performance characteristics. If we could show that groups composed of youngsters taken from an ongoing classroom with different rankings on Perceived Academic Ability would participate on the new task in relation to their relative ranking in their classroom, we might have some confidence that the explanation provided by this theory was suitable for the classroom situation.

If Perceived Academic Ability functions like General Performance Characteristics it helps account for patterns of consistent low achievement across a variety of subjects seen especially in students of low social status. More important, we could apply what we have learned about the manipulation of expectations stemming from status characteristics to alter general expectations for competence or incompetence in the classroom. Several studies have shown

that general expectations stemming from status characteristics may be offset by having the low status person demonstrate competence on two new interrelated performances (Freese & Cohen, 1972; Cohen, Roper et al, 1972). The work with race as a status characteristic has been called Expectation Training. The use of this theoretical framework to treat expectations would call for the treatment of the expectations held by the low achiever, the expectations for the low achiever's performance held by other students and by the teacher.

Another way to short-circuit the self-fulfilling prophecy which may be derived from this framework is a direct attack on the belief system of the teachers and the students. We may modify the belief that there is a single human ability dimension ranging from "smart" to "dumb," so that poor performance in a few schoolroom tasks is felt to imply that a person has a low general ability.

We could also interfere with the evaluation process by changing the reward structure. Individuals could be evaluated on their progress they are making compared to their own baseline rather than a competitive reward structure which defines the situation as necessarily having students with high grades and high achievement and students with low grades and low achievement.

Finally we could vary the tasks in the classroom, so students could display a much broader range of abilities to

each other than traditional verbal academic ability.

In review, this experiment is designed to examine the operation of expectations based on Perceived Academic Ability in classrooms in a new task fitting the specifications of the theory of Status Characteristics, General Performance Characteristics and Expectation States. If it can be shown that expectations for Perceived Academic Ability held by the student himself and by classmates for the student will influence a performance on a new non-academic task, we will have some confidence in designing classroom experiments based on this theory. We will have identified Perceived Academic Ability as a good instance of a General Performance Characteristic.

The hypotheses to be tested are derived from the theory: Group members who have higher rank on the General Performance Characteristic will be more active on a new task (irrelevant to previous performance) than will those who have a lower rank on the General Performance Characteristic.

The Task: Simulation Game

The criterion task for the proposed experiment has especial interest because it was an adaptation of a typical classroom simulation game, Seal Hunting. The use of a simulation game extends the "field setting" theme of this study by providing an authentic situation in which generalized academic performance characteristics may operate. To

date, research in Performance Characteristic Theory has been confined to laboratory settings where conditions and tasks have been strictly controlled. This study is aimed at the discontinuity between the "exact but synthetic" laboratory setting and the "less controlled but natural" setting of the classroom. Though not the primary interest of this study, a secondary concern will be to study the operation of a simulation game in the light of Performance Characteristic Theory.

By definition, simulation games closely resemble the theoretical scope conditions believed necessary to activate a General performance Characteristic. Simulation games usually involve group decision-making. They are a recent innovation and, as a result, many schools have not yet introduced them to their students. Upon their introduction to a simulation game, it is likely that children would not identify a specific performance characteristic with expected performance in the game. Rather, we suggest that the prevailing power and prestige order of students along the performance characteristic of Perceived Academic Ability may come into play and affect the differential performance of players of the game.

Some writers have claimed that the powerful, entertaining, relevant and motivating qualities of problem-solving simulations offer a rare chance for slow and bright children to work together effectively (Ingraham, 1967). From the point of view of Performance Characteristic Theory, this claim must be considered non sequitur until it can be tested empirically.

II

HYPOTHESIS AND DESIGN

The first objective of this study was to determine to what extent performance expectations, stemming from the General Performance Characteristic of Perceived Academic Ability (PAA), diffuse into small groups playing a non-academic and non-intellectual simulation game, thereby affecting the observed power and prestige order emergent in the group task.

The second objective of the study was to determine if one's task related performance in the group task is set at an <u>absolute</u> level, depending on his assigned position on the General Performance Characteristic, or whether the effects of one's academic status are <u>relative</u> to the particular positions on the performance characteristic of others in the group.

Briefly, the general design of the experiment was as follows: There were two conditions referred to as A and B. Condition A was made up of 20 four-man groups composed of boys ranked High and Medium on Perceived Academic Ability. There were two boys of each rank in each group. In Condition B, the four-man groups were composed of two members of High Perceived Academic Ability and two of Low Perceived Academic Ability. Perceived Academic Ability is an indicator of the ranking of the individual on the General Performance Characteristic. The groups play a game of Seal Hunting which is designed to meet the theoretical specifications of sufficient conditions for the activation of expectations flowing

from the General Performance Characteristic. The dependent variable is the power and prestige order emerging as the groups play the game. This power and prestige order is indicated by the amount of task-related verbal initiation made by each member of a group. Two hypotheses were tested:

1a. Holding constant social power, High *Ps.* subjects will have higher initiation rates than Medium *Ps.* subjects when they work together in a group. (Condition *A*)

1b. Holding constant social power, Medium *Ps.* subjects will have higher initiation rates than Low *Ps.* subjects when they work together in a group. (Condition *B*)

2. Subjects with Medium *Ps.* in Condition *B* will have higher initiation rates than subjects with Medium *Ps.* in Condition *A*.

PROCEDURE

Briefly, the steps of the procedure were as follows: First, boys in junior high school social studies classrooms rank-ordered themselves and their male classmates along the separate continua of Perceived Academic Ability and Social Power.

Second, four-man groups, composed of boys from the same classes, were selected from pre-existing classrooms and assigned to the two conditions. Two High *Ps.* boys and two Medium *Ps.* boys were assigned to groups for

Condition A. Two Medium PAA boys and two Low PAA boys were assigned to groups for Condition B.

Third, each four-man group played a new simulation game, Seal Hunting. The game lasted about 30 minutes. Each session was video-taped and analyzed to determine the power and prestige order as measured by the number of task-related speeches each member contributed to a group.

Selection of Subjects

The subjects were 156 Caucasian seventh and eighth grade boys from 35 different social studies classes in two middle class suburban schools. By controlling on race, social class, sex and age the effects of diffuse status characteristics were minimized. Any boy designated by his records or by his teacher as being mentally retarded was not included in the sample.

Subjects were ranked on the General Performance Characteristic of Perceived Academic Ability; within each class, members rank ordered themselves in terms of "school ability." (Brookover, 1962, 1965 and 1967). Each subject was shown a list of his male classmates and then asked to list the top three and the bottom three boys in terms of "school ability." The combined ratings for each boy were tabulated in rank order form and the top (High PAA) and bottom (Low PAA) three boys in each class were identified. Boys receiving little or no mention from their classmates were

identified as "medium" (Medium Pwr). The results of this measurement revealed an extremely high degree of agreement between students on the relative ranking of individuals as well as agreement between students and teachers.

The use of the theory of General Performance Characteristics required (1) that subjects have no specific experience with this new task and (2) that the effects of social power as opposed to achievement or performance status be controlled. Controlling the first factor was relatively easy; neither school had used simulation games.

Controlling for social power, on the other hand, was more complicated. For purposes of this study power is defined as ability to influence others derived from sources other than academic competence. Gold (1958) identified 17 sources of social power among elementary school children; ranking was determined by one's position on any one or more of these dimensions. Lippett and associates (1958) measured power along five dimensions and found that the combined index correlated highly with the single criterion of projected group influence "Who is best at getting others to do what he wants them to do?" In a later study Lippett and Gold (1959) reduced the index to four dimensions, the first three (expertness, affect, coercibility) independently and significantly correlating with the fourth (social power).

Since social power appears to be a unidimensional characteristic despite the fact that it may be derived from a

variety of sources we simply asked the subjects to identify the top and bottom three members in their class: "Who in your class is most able to get others to do things?" and "Who in your class is least able to get others to do things?" Rankings for social power were determined in the same way as for Perceived Academic Ability. The research design called for the use of boys who had been ranked as High, Medium or Low on Perceived Academic Ability, but who would be relatively equal in social power. Therefore, only subjects who had been ranked as Medium on the social power dimension were used in the study.

Steps were taken to guarantee the correspondence between self-attributed and other-attributed status on the two sociometric measures. The decision to use subjects who showed general agreement between the ratings they gave themselves and the average rating given them by their peers is based upon previous research findings which indicate that both sources (self and other), when considered separately, are significant determiners of individual behavior (Brookover, 1962, 1965 and 1967; Mischel and Staub, 1965).
Cases where there was a wide disparity between an individual's self rating and his average rating resulted in the elimination of that person from the sample. Parenthetically, we found surprisingly few cases where a child's self rating disagreed considerably with the average rating given him by his peers (about 30 out of over 600 cases).

Criterion Task: Seal Hunting

Under the two conditions of the experiment, each four-man team participated in the simulation game, Seal Hunting, which was developed under the direction of Jerome S. Bruner of Harvard.

Participants in Seal Hunting play the role of Eskimos attempting to survive in a hostile environment by catching seals. The game is a board game which simulates the chance and cognitive elements that determine success at hunting seals. The game requires a series of decisions as to which ice holes on the board should be "harpooned" in search of seals. For the experiment, the game was modified from its published form to make it a cooperative problem-solving task. In pretests, the modified version of the game demonstrated that it met the conditions which the theory specifies as sufficient for the activation of General Performance Characteristics in new task situations.

An important requirement of the theory is that the criterion task have decisions activating differential evaluation processes among participants. To accomplish this, the interdependence of players has been increased by restricting, to group decisions, those choices of where to hunt seals, which ice holes to harpoon, and by establishing group goals of survival and the attainment of the highest team score among all teams participating. It was believed that these constraints, in addition to the instructions that some

strategies are distinctly better than others, would produce a situation where players would be forced to evaluate each other's competence when arriving at the best possible strategy.

The game is sufficiently different from previous classroom experiences that specific expectations for competence derived from previous classroom performance are likely to be irrelevant. The non-academic nature of the game is critical. The nature of the game, hunting seals, and the simplicity of its rules were deemed sufficient to guard against it being recognized as an intellectual or academic group task. Quizzes on the rules given during the pre-tests showed that individual players had no difficulty understanding the rules.

The procedural steps within the basic design of this experiment are diagrammed in Figure 1.

MEASUREMENT PROCEDURES

Treatment of Data from the Game

Videotape recordings were made of each team as it played the game. The recording made of each team was then played back to a trained observer who scored the task-related verbal acts of each player. The scoring system, a modified Bales scheme, categorizes task acts into four types: performance outputs, action opportunities, positive evaluations and negative evaluations. For the purpose of this study, the four types of task acts were combined into one overall index of gross initiation.

The observers were three graduate students trained in the use of the scoring system. At the end of the training period each observer independently scored the same practice tape. Total initiation scores for subjects in the practice tape were compared across observers and analyzed according to a chi-square statistic. The observer's data yeilded a chi square of 2.26, df = 6. The probability that this difference occurred by chance is estimated to be above .85. Throughout the actual scoring period, every fourth group was independently scored by all three observers. These reliability checks all yielded probabilities above.75 that the disagreement between observers was due to chance.

RESULTS

The logic of the design and the hypotheses calls for the following comparisons:

1. Within each condition and within each group, the initiation rate of subjects with a higher rank on PAA must be compared with the initiation rate of members with a lower rank.
2. A between-Condition comparison must be made of medium PAA subjects who are relatively low ranking in Condition A and high ranking in Condition B.

These comparisons will be made using the rank order of members within a group and the percentage of task acts contributed by each subject to his group.

Finally a comparison of the total number of acts in Conditions A and B will be made in order to test the similarity of output. In order to show that initiation rates are a function of one's relative standing on PAA rather than a function of absolute behavioral differences of people having different ranks on PAA, it must be shown that the groups in the two conditions are equally active at the task.

Rank Order Analysis of Initiation

There were 39 four-man groups, 20 in Condition A and 19 in Condition B. The rank order analysis confirmed the hypotheses concerning the association of relative initiation rates with relative standing on PAA.

The four players in each group were rank ordered on relative number of task-related acts contributed to the group. The player who contributed the most acts received rank one; the player who contributed the second most acts was ranked second; etc. By combining all groups within each condition, one can determine the frequency and probability of a group in a given condition having a member of a given status holding top rank.

In Condition A, High PAA subjects are much more likely than Medium PAA subjects to hold the top rank (in 17 out of 20 groups). Similarly in Condition B the Medium PAA subjects are much more likely to hold the top rank than Low PAA subjects (15 out of 19 groups). Furthermore, comparison of the rank position of the Medium PAA subjects between conditions shows that they are much more likely to hold top rnak in Condition B than in Condition A (15 out of 19 groups vs.. 3 out of 20 groups). See Table 1.

Analysis of Percentage of Acts Contributed by Members

Rank order comparisons can be misleading when the actual differences between subjects' performances are narrow. In order to examine the size of the discrepancy between players on initiation rate, the percentage of all task-related acts contributed by each group member to the group total was calculated. Summing these individual percentages across groups and dividing by their

number gives the mean percentage of acts contributed by given types of subjects within each condition. (See Table 2)

Table 2 presents the means of the percentages of acts initiated by type of subject within each condition. Each pair of the same level PAA subjects was ranked and divided into "High Active" and "Low Active" depending on their relative standing on task initiation. The column under "Mean Percentage of Acts," gives the average percentage of initiation contributed by the subjects for each activity category for each value of PAA by condition. For example, the average percentage of acts contributed by more active High PAA subjects in Condition A (35.1%) is considerably higher than that contributed by less active High PAA subjects (22.19%).

The data provided by Table 2 allows another test of the hypothesis that subjects with higher rank on PAA will contribute more task related acts than their teammates who have a lower PAA. In Condition A for the more active subjects, High PAA players contributed a mean percentage of 35.1% of all group acts while Medium PAA subjects contributed, on the average, 25.9% of their groups' total. Similarly, for less active subjects, High PAA players contributed a higher percentage of acts (22.2%) than did the Medium PAA subjects (15.4%).

In Condition B it is also true that the mean percentage of acts contributed by the subjects with relatively

higher PAA is greater than that contributed by those with relatively lower PAA, for each activity category. Medium PAA players contributed on the average ~~22%~~ 27.1 of the teams' acts while their lower status teammates (Low PAA) accounted for ~~17.6%~~ 22.5 of the acts.

In each condition, when both activity categories are combined, the average percentage of acts contributed by each higher status player to his team's effort is larger than the average contributed by a lower status player (~~22.9~~ 29.1% for High PAA subjects vs. 20.6% for Medium PAA subjects in Condition A and 27.6% for Medium PAA subjects vs. 22.5% for Low PAA subjects in Condition B).

Comparison of Mean Percentages of Acts between Conditions

The means of the percentages of acts initiated by different status levels for differing activity rates allow several statistical comparisons to be made. Figure 2 makes these comparisons clear.

This graphic presentation shows the differences between the performance of the more and less active Medium PAA subjects in Condition A compared to Condition B. In addition, the graph shows the similarity between the performance of the relatively high status subjects in Condition A with those of relatively high status in Condition B. The same pattern of similarity in activity rates between conditions can be noted for those of relatively low status position.

Controlling for the other factors which might affect performance, it was hypothesized that the difference in the behavior of Medium PAA players between Conditions A and B could be accounted for by the differential reaction of this type of subject to the perceived academic status of his teammates. It was expected that when Medium PAA subjects were the low status players in Condition A teams, their average percentage of acts contributed could be significantly less than the average for Medium PAA subjects in Condition B where they were the high status players.

A randomized test for two independent samples (t-test) was used to compare the difference between conditions in the mean percentage of all acts initiated in each group by Medium PAA subjects.[1] Comparisons were made for all the Medium PAA subjects in each condition and for Medium PAA subjects separated by activity rates. Table 3 reports the results of these statistical tests.

These statistical tests illustrate what was pointed out in Figure 1: Regardless of activity rates, subjects of Medium PAA were significantly more active when teamed with boys of Low PAA than when teamed with boys of High PAA.

In review, the results of the analysis illustrate the effect on initiation rates of the rank on PAA relative to the rank of other group members.

---

[1] S. Siegal, Non-Parametric Statistics for the Behavioral Sciences, New York: McGraw-Hill, 1956, p. 152.

Similarity Between Conditions

Before we can accept the above conclusions about the effects of generalized academic performance expectations, it must be shown that differences in subjects' performances between conditions were not a function of qualitative differences between the two conditions.

The simplest way to show the similarity between Conditions A and B is to compare the average team performance between conditions. Table 4 shows the results of this comparison using a randomized test for two independent samples (t-test).

There is no significant difference between the average number of acts initiated by teams in Condition A (385.4) and teams in Condition B (376.3). Since the average number of initiated acts for teams in each condition is nearly the same, differences in mean percentage of acts contributed by given types of subjects represent absolute as well as relative differences in performance.

The finding that teams in each condition performed similarly is further strengthened by the data shown in Table 5. Disregarding the perceived academic ability of subjects, a statistic comparison was made between conditions of the mean percentage of total group acts for each rank order position. In terms of mean percentages of total group acts, there were no differences in the way subjects in a particular rank order behaved in Condition A when compared

with similarly ranked subjects in Condition B.

Finally, the similarity of conditions is supported by data in Table 6. Holding the level of activity and academic status constant, comparisons were made across conditions. There are no significant differences in mean percentages of total group acts between subjects who held the same academic status position relative to their teammates on teams in Condition A as compared to those in Condition B. For example, the mean percentage of team acts initiated by the more active High PAA subjects in Condition A did not differ significantly from that of the more active Medium PAA subjects who held the higher state of the general performance characteristic in Condition B.

The finding of no significant differences in percentage of acts initiated is consistent for all other comparisons made between conditions of players holding relatively higher and lower PAA rank within their teams.

INTERPRETATION

The results are consistent with the hypothesis that when diffuse status characteristics and social power are controlled, performance expectations, stemming from the General Performance Characteristic of Perceived Academic Ability, act like a diffuse status characteristic in affecting the emergent power and prestige order in a four-man group of junior high school boys. In this case the boys

were engaged in a new group task considerably different from tasks where Perceived Academic Ability would function as a reasonable predictor of performance.

Of especial significance is the comparison showing the students of Medium PAA were significantly more active when combined with Low PAA subjects than when they worked with High PAA subjects. This finding implies that a person with this rank will behave significantly differently, depending on the state of his partners. It appears that performance expectations do not function in such a way that assignment on the General Performance Characteristic indicates an absolute level of performance in a new setting (as if GPC were a characteristic of the individual), but indicates a level of performance relative to the status comparisons in the work group. This interpretation is further strengthened by the finding of no significant differences in activity levels between the two conditions.

In its adaptation for this study, Seal Hunting proved to be a strictly non-intellectual activity. Regardless of condition, the trained observers were struck by the general lack of intellectual discussion that might have aided teams in the capture of seals and the tendency of players to ignore the useful information made available to them in the game's instructions. Typically, teams in both conditions played the simulation game without the use of any particular rationale. This observation, coupled with the

data showing no significant difference in the performance of teams between conditions, strengthens the contention that those subjects who contributed more of their team's acts did so not because the game favored their higher IQ score. What we have is a new task which is irrelevant to commonly accepted classroom skills or academic intelligence; yet those members of the group who are higher in perceived academic ability are far more active in the game than those who are lower.

The irrelevance of the game to classroom performance expectations was further increased by the conditions under which it was played in the experiment. Each four-man group left its social studies classroom and reported as an individual team to a utility room, normally not used for instructional purposes. The room had none of the trappings usually associated with a typical junior high school classroom. Except for the video recording apparatus, which was kept at a distance, the only other furnishings were four chairs and a game table. The players were introduced to the game by a tape recording which explained that a company, which made games for children their age, simply wanted to know if the game were engaging enough to warrant its production and marketing. The boys were told by the experimenter to relax and to enjoy this chance to "skip class." It was believed that the mere fact that the game was run in the school setting, even though not in a classroom, would be

enough to activate expectations for performance based on the General Performance Characteristics of Perceived Academic Ability. The results of the experiment strongly support this assumption.

IMPLICATIONS

The most impressive and perhaps distressing finding for educators is that generalized academic performance expectations act like a diffuse status characteristic or a general performance characteristic in affecting the performance (power and prestige order) of children engaged in a new school task--a task which in this case was irrelevant to normal academic skills and classroom routine. Here is a General Performance Characteristic created in the independent task setting of the classroom which behaves just as if it had been constructed by an experimenter under laboratory control. Considering the length or time and opportunity these children had to learn about each other, the fact that they could behave as if the only thing they knew about each other was Perceived Academic Ability is fairly shocking. Perhaps this is because conventional classrooms in the junior high school do not provide too great a variety of experience where children can see the many different talents and abilities each possesses.

Implications for classroom performance are not pleasant to contemplate. One may infer that exciting new

curricular tasks such as game simulation, no matter how different from conventional curriculum, are very likely to be affected by general performance expectations for competence and incompetence built up through classroom interaction. The contention, made earlier, that simulation games may offer a rare chance for the "bright" and the "slow" student to work together effectively is simply not supported. Thus, we cannot look to purely curricular solutions to the problem of the inactive, failing student.

Another important finding of the study was that the junior high school boys in the sample were clearly able to differentiate classmates on at least three levels of Perceived Academic Ability (High, Medium and Low). An extremely high degree of agreement was found between students, as well as agreement between students and teachers, on the relative ranking of individuals. The results of the experiment indicate that these three relative rank orders represent actual differentials in performance expectations which, when activated, affect corresponding differentials in individual performance at new group tasks. Previous literature on the self-fulfilling prophecy stresses the importance of expectations held by the teacher. This study shows that expectations held by other students and by the person himself can trigger a self-fulfilling prophecy.

To be sure, the middle class and suburban character of the subjects of the experiment may limit generalization

about the effect of generalized academic performance expectations to similar type populations. The sources of classroom status may be quite different in inter-urban or rural populations. Additionally, academic reputation may not be as significant a source of status in schools less traditional in curriculum and classroom organization. The two junior high schools chosen for the experiment were ones in which the General Performance Characteristic of Perceived Academic Ability would most likely operate. In both schools a large majority (70-80%) of the pupils eventually go on to college. One of the schools had a social studies class especially reserved for academically advanced students. The social studies classes, from which subjects were drawn, appeared to be traditionally academic content-oriented and were operated as closed classrooms.

Granted that generalized academic performance expectations do affect performance at new school tasks, what featuresof the classroom can be inferred as responsible for the emergence of general expectations for academic competence or incompetence? We have chosen the evaluation process in the classroom as one of the critical pre-conditions for the emergence of the General Performance Characteristic of Perceived Academic Ability. There are several ways in which classroom evaluation procedures make it generally known what level of performance is to be expected

from each child. One way is the much-discussed method of ability grouping which publicly labels the child according to specific expectations for his performance in reading or math or is sometimes used as a general label as in tracking systems.

Studies of ability grouping indicate that the practice has consequences far beyond academic achievement. Teachers in such systems tend to develop more rigid opinions concerning individual differences in children; they divide children into types--the bright, the average and the dull. Moreover, these are thought of as enduring immutable, unchangeable, unassailable attributes of the children (Lackman and Secord, 1966). A study of the parallel practice in British Schools finds that streamed schools overemphasize competition while in unstreamed schools, children were more cooperative and helpful toward each other (Jackson, 1964).

A third source of expectations is the general performance expectations created by competition within the classroom. In some extreme cases of low ability grouping or very low status classroom, this competitive element is probably absent. Even if the teacher has these norms in mind, there are enough children who do not perceive this dimension of school "smartness" as relevant or important so that the process essentially fails to operate. But in typical classroom settings, the competitive model underlying so many classroom tasks provokes frequent social comparisons along

a relatively simple dimension of perceived ability. Students making these comparisons are not thinking of a set of specific and uncorrelated competencies but are making a specific success or failure as continuing evidence for one's place along the underlying general dimension of competence.

The identification of sources of performance expectations in the classroom suggests a way that the currently narrow notion of perceived ability may be broadened. If it can be determined that the competitive tone of most classrooms is a source of performance expectations, what strategies can be suggested to lessen the impact of the competitive model? One strategy might be to decrease the role of the teacher as an evaluator, making him more of a resource person. The elimination of both grades and public recitations might also be recommended. The strategy may call for the development of educational tasks which involve many different kinds of skills. Within these new tasks, allowance would be made for increased interaction opportunities between individuals so that they may come to recognize that both they and their classmates possess many different talents and abilities.

FIGURE 1

RESEARCH DESIGN

STEP 1

Selection of Subjects

Controls:
1. Age
2. Sex
3. Race
4. Social class
5. Prior game experience
6. Social power

STEP 2

Administer Sociometric Device

Subjects rank ordered on:
1. Perceived academic ability
2. Social power

STEP 3

Assignment to Conditions

Condition A	Condition B
20 four-man teams	19 four-man teams
2 boys (High PAA)	2 boys (Medium PAA)
2 boys (Medium PAA)	2 boys (Low PAA)

All subjects ranked Medium on Social Power

STEP 4

Play Game

Seal Hunting

FIGURE 2

MEAN PERCENTAGE OF ACTS INITIATED BY MORE AND LESS ACTIVE MEMBER OF EACH STATUS PAIR: BY CONDITION

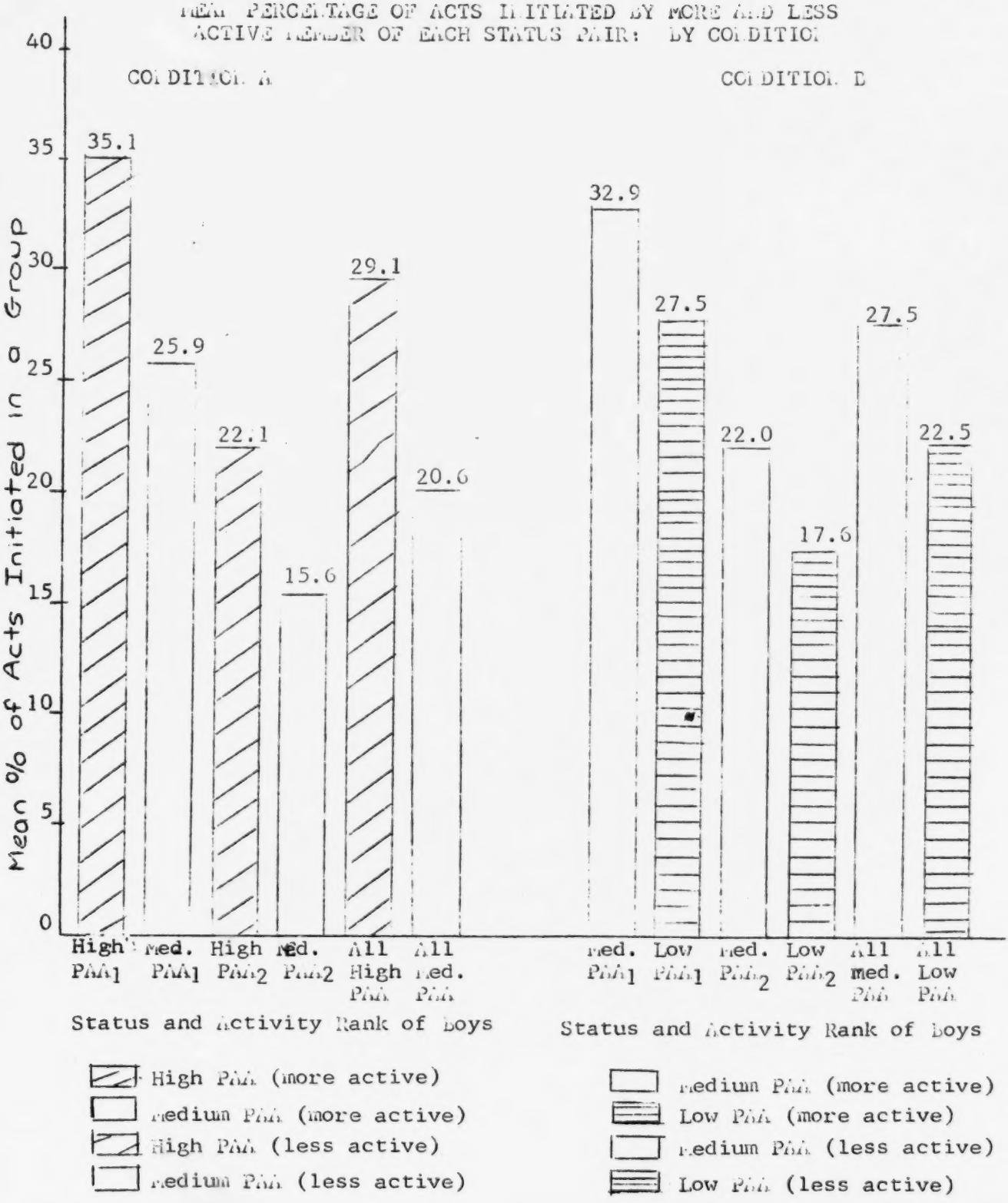

TABLE 1

PROBABILITY OF EACH RANK BEING HELD BY A SUBJECT AT LEVEL OF PAA AND CONDITION

CONDITION A

Rank in Group	High PAA		Medium PAA	
	Number of Groups	(Probability)	Number of Groups	(Probability)
1 or 1.5	17	(81)	4	(19)
2 or 2.5	9	(45)	11	(55)
3 or 3.5	10	(50)	10	(50)
4	4	(21)	15	(79)

CONDITION B

Rank in Group	Medium PAA		Low PAA	
	Number of Groups	(Probability)	Number of Groups	(Probability)
1 or 1.5	15	(79)	4	(21)
2 or 2.5	10	(53)	9	(47)
3 or 3.5	7	(37)	12	(63)
4	6	(32)	13	(68)

*Number of groups does not add to 20 because of ties.

TABLE 2

MEAN PERCENTAGE OF TOTAL GROUP ACTS INITIATED BY EACH LEVEL OF PERCEIVED ACADEMIC ABILITY, FOR TWO CONDITIONS: HOLDING THE LEVEL OF ACTIVITY CONSTANT

Activity Level	Perceived Academic Ability	Mean Percentage of Acts	N of Groups	Total % of Group Acts Contributed by Each Status Pair
		Condition A		
High	High PAA	35.10	20	
	Medium PAA	25.94	20	
Low	High PAA	22.19	20	
	Medium PAA	15.35	20	
Combined Activity Levels	Mean of High PAA	29.14	20	58.28
	Mean of Medium PAA	20.63	20	41.29
		Condition B		
High	Medium PAA	32.85	19	
	Low PAA	27.46	19	
Low	Medium PAA	22.04	19	
	Low PAA	17.59	19	
Combined Activity Levels	Mean of Medium PAA	27.45	19	54.89
	Mean of Low PAA	22.52	19	45.05

N of subjects in Condition A = 80
N of subjects in Condition B = ~~80~~ 76

TABLE 3

RANDOMIZATION TEST FOR TWO INDEPENDENT SAMPLES: MEAN PERCENTAGES OF GROUP INITIATION CONTRIBUTED BY MORE AND LESS ACTIVE SUBJECTS OF MEDIUM PERCEIVED ACADEMIC ABILITY IN THE TWO CONDITIONS

	t score		
Conditions Compared	More Active Medium PAA	Less Active Medium PAA	All Medium PAA
A X B	-3.595*	-3.447*	-3.740*

Note: Scores are based on 20 more active and 20 less active medium PAAs in Condition A; and 19 more and 19 less active medium PAAs in Condition B.

*$p < .005$ for one tailed test

TABLE 4

RANDOMIZATION TEST FOR TWO INDEPENDENT SAMPLES: AVERAGE NUMBER OF ACTS INITIATED BY TEAMS IN EACH CONDITION

	Condition A	Condition B	t-score A X B
Average Number of Acts	385.4	376.3	0.254 (n.s.)

Note: Scores are based on 20 teams in Condition A and 19 teams in Condition B.

TABLE 5

RANDOMIZED TEST FOR TWO INDEPENDENT SAMPLES: MEAN PERCENTAGES OF TOTAL GROUP ACTS CONTRIBUTED BY EACH RANK ORDER POSITION, REGARDLESS OF PERCEIVED ACADEMIC STATUS, ACROSS CONDITIONS

Comparisons	Average % of Acts Initiated by Each Rank			t-score	Level of Significant Differences
	Condition A		Condition B		
1	Rank 1 (36.1)	vs	Rank 1 (35.2)	-0.094	n.s.
2	Rank 2 (27.6)	vs	Rank 2 (27.3)	-0.481	n.s.
3	Rank 3 (21.9)	vs	Rank 3 (22.0)	-0.143	n.s.
4	Rank 4 (14.3)	vs	Rank 4 (15.4)	-0.773	n.s.

Note: There are 20 subjects in each rank in Condition A; and 19 subjects in each rank in Condition B.

TABLE 6

RANDOMIZED TEST FOR TWO INDEPENDENT SAMPLES: MEAN PERCENTAGE OF TOTAL GROUP ACTS INITIATED BY EACH LEVEL OF PERCEIVED ACADEMIC ABILITY, ACROSS CONDITIONS, HOLDING THE LEVEL OF ACTIVITY CONSTANT

Comparisons	Activity Level	Groups		t-score	Level of Significant Differences
		Condition A	Condition B		
1	High	High PAA (35.1)	vs Medium PAA (32.9)	-1.430	n.s.
2		Medium PAA (25.9)	vs Low PAA (27.5)	-0.705	n.s.
3	Low	High PAA (22.2)	vs Medium PAA (22.0)	-0.063	n.s.
4		Medium PAA (15.4)	vs Low PAA (17.6)	-1.294	n.s.
5		Average of all High PAA (29.1)	vs Average of all Medium PAA (27.5)	-0.943	n.s.
6		Average of all Medium PAA (20.6)	vs Average of all Low PAA (22.5)	-1.049	n.s.

BIBLIOGRAPHY

1. Backman, C.W. and P.F. Secord. A Social Psychological View of Education. New York: Harcourt, Brace and World, Inc., 1968.

2. Berger, J., B.P. Cohen and M. Zelditch. "Status Characteristics and Expectations States," in J. Berger, et. al. (eds.) Sociological Theories in Progress. Boston: Houghton Mifflin, 1964, 29-46.

   _____, and T. L. Conner. "Performance Expectations and Behavior in Small Groups," Acta Sociologica, 14: 186-198, 1969.

3. Brookover, Wilbur, et. al. "Self-Concept of Ability and School Achievement," Office of Education, Cooperative Research Project No. 845, Bureau of Educational Research Services, College of Education, Michigan State University, 1962.

4. _____, J. Le Pere, D. Hamachek, T. Shaila and E. Erickson. "Self-Concept of Ability and School Achievement, II: Improving Academic Achievement Through Students' Self-Concept Enhancement," Office of Education, Cooperative Research Project No. 1636, Bureau of Educational Research Services, College of Education, Michigan State University, 1965,

5. _____, et. al. "Self-Concept of Ability and School Achievement, III," Office of Education, Cooperative Research Project 2831, Bureau of Educational Research Services, College of Education, Michigan State University, 1967.

6. Brophy, J. and T.L. Good. "Teachers' Communication of Differential Expectations for Children's Classroom Performance: Some Behavioral Data." Journal of Educational Psychology, 1970, 61, 365-374.

7. Cohen, E.G. "Interracial Interaction Disability: a Problem for Integrated Education," Urban Education, 1975, 5, 336-356.

   _____, et. al.. "Expectation Training I: Altering Effects of a Racial Status Characteristic," Technical Report No. 2, School of Education, Stanford University, 1970.

8. _____, and Susan S. Roper. "Expectation Training II: Modification of Interracial Interaction Disability," Technical Report No. 7, School of Education, Stanford University, 1971.

   Deal, Terry E. "Modifying the Effects of Age Within Decision-Making Groups," Technical Report No. 4, School of Education, Stanford University, 1971.

9. Freese, Lee. "The Generalization of Specific Performance Expectations," Unpublished Ph.D. dissertation, Stanford University, 1970.

10. _____, and B.P. Cohen. "Eliminating the Effects of the Diffuse Status Characteristic," Technical Report, Laboratory of Social Research, Stanford University, 1971.

11. Gold, M. "Power in the Classroom," Sociometry, 25: 50-60, 1958.

12. Ingraham, L.W. "Teachers, Computers and Games: Innovation in the Social Studies," Social Education, 31: 51-53, 1967.

13. Jackson, Brian, Streaming: An Education System in Miniature, London: Routeledge and Kagan Paul, 1964.

14. _____, and M. Gold. "Classroom Social Structure as a Mental Health Problem," Journal of Social Issues, 15: 40-58, 1959.

15. Mischel, W. and E. Staub. "Effects of Expectancy on Working and Waiting for Larger Rewards," Journal of Personality and Social Psychology, 2: 625-633, 1965.

16. Rosenthal, R. and L. Jacobson. Pygmalion in the Classroom, New York: Holt, Rinehart and Winston, 1968.